Structure of the B-DNA Oligomers d(CGCTAGCG) and d(CGCTCTAGAGCG) in New Crystal Forms†

Valya Tereshko,‡ Lourdes Urpí,‡ Lucy Malinina,§ Tam Huynh-Dinh,|| and Juan A. Subirana*,‡

Departament d'Enginyeria Química, Universitat Politècnica de Catalunya, Avenida Diagonal 647, E-08028 Barcelona, Spain, Engelhardt Institute of Molecular Biology RAS, Vavilova 32, 117984 Russia, and Unité de Chimie Organique, Institut Pasteur, 32 Rue du Dr. Roux, Paris 75724, France

Received May 1, 1996; Revised Manuscript Received June 26, 1996⊗

ABSTRACT: We present the structure of the dodecamer CGCTCTAGAGCG and the related octamer CGCTAGCG, both in the B form, determined by single crystal X-ray diffraction. Two different crystal forms of the octamer have been obtained, with either three or four duplexes in the asymmetric unit. The dodecamer crystallizes in the $P2_1$ space group with two duplexes in the asymmetric unit. Very few such structures have been previously reported, while the octamer structure is the first one determined with three duplexes in the asymmetric unit. It is also the first octamer with standard Watson−Crick base pairs to be crystallized in the B form. The crystal structure is stabilized in both cases by interactions between the guanines in the two terminal base pairs of each duplex. This interaction is similar to that found in most dodecamers which have been previously studied, but here it is found in a new unit cell (for the dodecamer) and in one octamer. In the dodecamer cytosine-stacking interactions between neighbor duplexes are also present. The two dodecamer duplexes in the asymmetric unit show different patterns of bending, while the octamer molecule has a rather straight helical axis. The results presented confirm the strong conformational variability of the TA pyrimidine−purine step and demonstrate a clear alternating structure for the $(CT/GA)_n$ sequence.

Since the pioneering work of Wing et al. (1980) a substantial number of oligonucleotide structures in the B form have been reported, either isolated or combined with proteins (Berman et al., 1992). Nevertheless, more sequences remain to be studied (Subirana et al., 1995). In the case of isolated oligonucleotides most of them are dodecamers and decamers which have been crystallized in very few space groups, with similar spatial organizations. The space group creates restrictions on the conformation of the crystallized oligonucleotides, and therefore, it is of interest to study new sequences in different crystal environments. In this paper we report the structure of two such oligonucleotides. A short report on their structures has been presented (Urpí et al., 1996) in which the conformation of the central CTAG sequence was compared with the same sequence in the *met* and *trp* operators.

The oligonucleotide structures presented here, besides their new sequence, offer some intrinsic unique features. We present the first octamer structure with standard Watson−Crick base pairs which crystallizes in the B form; most octamers previously crystallized have been found in the A form. The oligonucleotide crystals reported here make use of the guanine−guanine interactions between the two terminal base pairs of neighbor duplexes, typical of B-DNA dodecamers (Dickerson et al., 1987), but the crystal system and packing organization are quite different. In fact, no octamer starting with the CG sequence, which may favor the B form through specific end-to-end interactions, has been previously crystallized in the A form. Another feature of interest is that both oligonucleotides have been found to crystallize with several duplexes in the asymmetric unit. Very few oligonucleotide structures with more than one duplex in the asymmetric unit have been previously studied. This feature allows a comparison of the same duplex in different packing environments in the same crystal.

MATERIALS AND METHODS

Synthesis and Purification. The self-complementary oligonucleotides were synthesized on an automatic synthesizer by the phosphoramidite method and purified by gel filtration and reverse-phase HPLC. The ammonium salt was prepared by ion-exchange chromatography and used for crystallization.

Crystallization. Crystallization was carried out by the vapor diffusion technique. The crystallization solution for the dodecamer was a sitting drop with 0.37 mM duplex, 27 mM calcium chloride, 7.5 mM sodium cacodylate (pH 7), and 6% MPD (2-methyl-2,4-pentanediol). This solution was equilibrated against a reservoir with 20% MPD at the beginning and 25% MPD at the end. One crystal of $0.7 \times 0.3 \times 0.2$ mm^3 obtained after about 7 days was used for data collection.

With the octamer two crystal forms were obtained from the same sitting drop. Initial conditions were 0.5 mM duplex, 9.0 mM L-alanyl-L-arginine acetate, 6.6 mM sodium cacodylate (pH 6.2), 3.75 mM magnesium chloride, 0.34 mM spermine tetrahydrochloride, and 9.75% MPD. This solution was equilibrated against a reservoir with 37.5% MPD which

† This work was supported in part by Grants PB93-1067 from the DGICYT, GRQ93-3030 from the DGU, and INTAS-93-2755 and by CESCA for free computer time. V.T. was supported by a MEC fellowship.

* Correponding author. Tel: 34-3-4016688. Fax: 34-3-4017150. E-mail: subirana@eq.upc.es.

‡ Universitat Politècnica de Catalunya.

§ Engelhardt Institute of Molecular Biology RAS.

|| Unité de Chimie Organique, Institut Pasteur.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1996.

S0006-2960(96)01048-3 CCC: $12.00 © 1996 American Chemical Society

Table 1: Crystallographic Data and Refinement Statistics

	dodecamer	octamer
unit cell parameters		
a (Å)	35.05	24.77
b (Å)	68.91	41.52
c (Å)	26.15	115.89
β (deg)	91.16	
space group	$P2_1$	$P2_12_12_1$
asymmetric unit	2 duplexes, 47 water molecules, Ca^{2+} ion	3 duplexes, 34 water molecules
data collection statistics		
resolution (Å)	2.25	2.47
collected reflections	12447	31617
unique reflections	5276	3742
R_{sym} (%)	8.9	8.8
completeness (%)	88.9	80.6
refinement parameters		
resolution (Å)	8.0−2.25	8.0−2.47
obsd reflections ($F > 2\sigma(F)$)	4222	3309
completeness (%)		
overall	73.1	75.2
8−2.9 Å	88.7	75.5
2.9−2.7 Å	70.9	74.8
2.7−2.5 Å	61.4	74.5
2.5−2.25 Å	51.5	
R-value (%)	19.7	18.2
rms bond lengths (Å)	0.022	0.018
rms bond angles (deg)	4.1	3.7

was later increased to 42.5% MPD. At this point small crystals which had been obtained in very similar conditions were seeded, and thin needles appeared in 2 days. Since the crystals did not grow further, after 15 days they were dissolved by decreasing the MPD in the reservoir to 30%. Then the MPD concentration was again slowly increased during 2 months up to 50%. Five weeks later crystals of about $2 \times 0.2 \times 0.2$ mm^3 were obtained. With such crystals the data presented here were collected. In the same drop, without changing conditions, crystals with a different morphology appeared 8 months later. They belong to the $P6_1$ space group, with four duplexes in the asymmetric unit ($a = b = 48.7$ Å; $c = 115.9$ Å). Data were also collected from this second crystal form, but have only been studied in a preliminary fashion. Packing of the octamers is very similar to that described below in the $P2_12_12_1$ space group.

Data Collection. Precession photographs were taken in order to determine the space group. In the case of the dodecamer it was clear that the molecules were in the B form with their axis oriented along the *b* axis of the cell. In the case of the octamer neither the DNA form nor the molecular orientation could be ascertained from the precession photographs.

An Enraf−Nonius GX-21 rotating anode generator operating with Cu Kα radiation was used. In the case of the dodecamer the X-ray data were collected in a CAD-4 automatic diffractometer and were corrected for polarization, Lorentz, absorption, and decay. In the case of the octamer the X-ray data were collected in a Mar Research image plate. Two spectra with different crystal−detector distances were collected from one crystal. A third spectrum was obtained from another isomorphous crystal. All data were processed with the MarXDS program (Kabsch, 1988). The three independent data sets were merged with the MARSCALE program. The correlation factor between equivalent intensities was 97%. Crystallographic data for both structures are given in Table 1.

Structure Determination and Refinement of the Dodecamer. The dodecamer structure was solved by molecular replacement using the AMoRe program (Navaza, 1994). The unit cell volume and the precession pictures indicate that the asymmetric unit of the crystal contains two B-DNA duplexes. The starting model used to solve the structure was a standard B-DNA duplex created with the NAHELIX program (Westhof, 1984), but it was not possible to find a solution with such model. An appropiate solution was found when only the central decamer was used as a model. The decamer molecules were refined by simulated annealing with the X-PLOR program (Brunger, 1992), and it was found necessary to displace one molecule by one base pair (rotation = 36°; displacement = 3.4 Å). Further improvement of the model was achieved by adding the terminal base pairs and by rigid-body refinement with AMoRe, using 24 rigid bodies, i.e., the 24 base pairs of the two duplexes in the asymmetric unit. In this way the *R*-factor decreased from 51.4% to 41.9% in the resolution range 15−2.5 Å. At this stage we started refinement with the X-PLOR program using all reflections with $F > 2\sigma(F)$. In the first stages of refinement noncrystallographic restraints were applied in order to minimize the difference between both dodecamer duplexes. Rigid body refinement with X-PLOR was carried out by splitting the model into an increasing number of rigid groups, up to 138 (sugars, phosphates, and bases). The resolution range used was 8−3 Å and the *R*-factor decreased from 40.9% to 29.8%. Then simulated annealing molecular dynamics were applied in the resolution range 8−2.25 Å; the *R*-factor decreased to 23.95%. Finally, positional refinement (8−2.25 Å resolution range) and refinement of individual *B*-factors (6−2.25 Å resolution range) were carried out without any noncrystallographic retraints. In the final part of refinement all reflections with a small intensity ($F_o < 1.5$) were eliminated. Thus the observed reflections decreased from 5135 to 4222 as indicated in Table 1. One calcium ion and 47 water molecules were introduced after inspection of the ($2F_o - F_c$) and ($F_o - F_c$) maps. The final *R*-factor was 19.7% in the resolution range 8−2.25 Å. Coordinates have been deposited in the Nucleic Acid Database (Accession number BDL070). A stereoview of both dodecamers in the asymmetric unit is shown in Figure 1.

The ($2F_o - F_c$) map displayed very poor electron density for the first cytosine (C1) of both duplexes in the asymmetric unit. Therefore, this cytosine was not included in the model. Since there are strong guanine−guanine interactions and several duplexes are present in the neighborhood, the terminal cytosine 1 might have different alternative positions and appear disordered. An electron density peak was found among four phosphate groups and was interpreted as a calcium ion, as shown in Figure 2. The intermolecular distances calculated are rather small. If this peak was refined as a water molecule, the average oxygen−water distances increased to 2.55 Å, a value too small for a water molecule.

Structure Determination and Refinement of the Octamer. The octamer structure was also solved with the help of the AMoRe program. A solution was attempted with an octamer with either standard A- or B-DNA geometry, but no adequate result was obtained. Then another model was constructed from the dodecamer structure obtained as described above by adequate docking of the coordinates of the CG terminal base pairs and the central CTAG tetramer. With this model

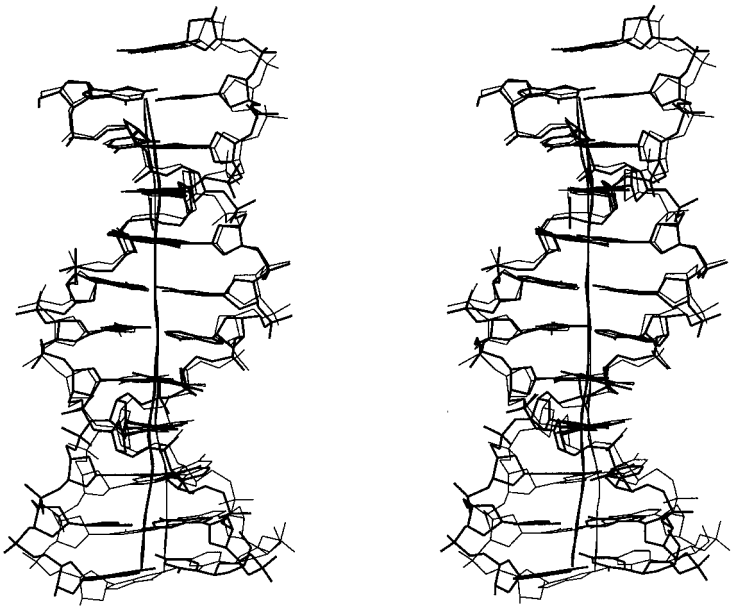

FIGURE 1: Stereo pair showing a superposition of the two dodecamer structures. Molecule D1 is represented with thick lines. The axes of the molecules, calculated with the program CURVES (Lavery & Sklenar, 1988), are also shown. The missing cytosine 1, discussed in the text, should be in the upper end.

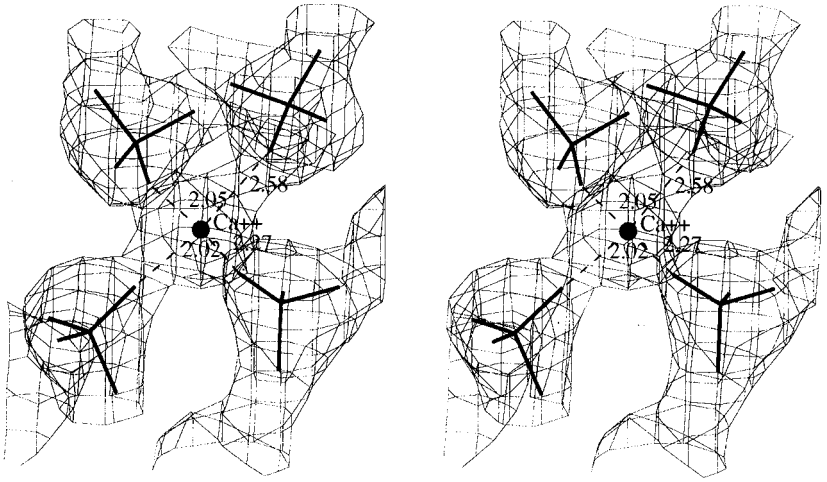

FIGURE 2: Stereo pair showing the electron density map around a calcium ion which interacts with the phosphate groups at the ends of four different duplexes in the crystal. The ion−oxygen distances calculated after refinement are indicated.

it was possible to find all three molecules in the asymmetric unit.

Once the molecules had been located, we used again a B-DNA duplex model with standard geometry for refinement. In the first steps pseudohexagonal symmetry was imposed, so that only one of the three molecules was considered as independent. Once the structure was refined in this way, the noncrystallographic restraints were removed and 34 water molecules were added stepwise. The same strategy of refinement used in the dodecamer case was applied for the octamer. Final refinement parameters are presented in Table 1. Coordinates have been deposited in the Nucleic Acid Database (Accession number BDH071).

Conformational Parameters. The conformational parameters given in this paper were calculated with the NEWHELIX 93 program, obtained from the Protein Data Bank. Some of them are given in Tables 2 and 3. A complete list is given as Supporting Information. In the case of the octamer the three molecules are very similar, a result due in part to the method of refinement and the limitations of resolution. The values for molecule 1 are used in Table 2 and in the figures, since this molecule had the most conformational parameters approximately equal to the average value of the three molecules in the asymmetric unit.

Table 2: Selected Conformational Parameters of the Octamer d(CGCTAGCG)[a]

base	propeller twist (deg)	twist (deg)	rise (Å)	roll (deg)
C1	−1.9	38.6	3.14	−0.1
G2	−16.7	37.1	3.46	−0.5
C3	−8.3	27.6	4.02	3.2
T4	−10.1	48.9	3.00	−6.4
A5	−14.4	27.8	3.84	0.7
G6	−5.8	38.7	3.36	−1.1
C7	−15.2	38.3	3.70	3.4
G8	−5.4			

[a] The values given correspond to the first octamer, which are approximately equal to the average of the three molecules in the asymmetric unit.

It should be noted that the intrinsic error of many of the DNA conformational parameters is very large, as indicated by the comparison of different refinement methods which have been reported (Westhof, 1987; Hahn & Heinemann, 1993) for structures studied at a higher resolution. It appears that the twist parameter is the least sensitive to the refinement

Table 3: Selected Conformational Parameters of the Two Dodecamers d(CGCTCTAGAGCG) in the Asymmetric Unit (D1 and D2)[a]

	propeller twist (deg)		twist (deg)		rise (Å)		roll (deg)	
base	D1	D2	D1	D2	D1	D2	D1	D2
G2	−12.3	−6.4	33.8	33.3	3.51	3.52	1.8	0.6
C3	−12.9	−8.3	34.7	34.5	3.77	3.96	7.4	4.4
T4	−8.3	−9.4	33.6	38.7	3.69	3.49	5.2	2.0
C5	−3.6	−8.6	29.3	29.5	3.65	3.50	0.7	0.6
T6	−5.1	−7.7	45.2	43.4	3.57	3.47	−3.8	0.6
A7	−16.7	−15.7	32.3	30.6	3.54	3.61	1.1	5.7
G8	−12.4	−12.3	36.0	39.0	3.35	3.34	7.1	0.0
A9	−17.0	−11.1	36.4	30.2	3.70	3.99	6.6	14.6
G10	−14.7	−19.7	38.6	38.7	3.05	2.93	−0.8	−1.2
C11	−13.3	−13.2	37.8	41.8	3.69	3.36	4.8	4.3
G12	−23.9	−22.4						

[a] The first cytosine (C1) could not be located, as described in the text.

method and the resolution available (Hahn & Heinemann, 1993; Subirana et al., 1995), since it defines the overall position of the base pairs around the helical axis, a feature which is very clear even at low resolution. Most of the other conformational parameters calculated by NEWHELIX are interrelated as shown by Yanagi et al. (1991). The rise parameter (D_z) is additive for consecutive base steps, since it is a projection on the average helical axis. It gives an idea of the local (and average) pitch of the helix. Yanagi et al. (1991) introduced a new overall parameter, the Profile Sum, which has been used by some investigators. If we introduce the relation between rise, cup, and roll derived by these authors (Formula 4) in their formula 6 for the profile sum, we obtain the expression:

$$\text{profile sum} = (\omega - 36) - 32.5(D_z - 3.36) \qquad (1)$$

This expression shows that the profile sum corresponds to a linear combination of twist and rise. Therefore, it appears that the latter two parameters are sufficient in order to compare different structures.

RESULTS AND DISCUSSION

Octamer Packing. The packing mode of the d(CGCTAGCG) octamer duplexes along the crystallographic *c* axis is shown in Figure 3. The six duplexes are tilted with respect to this axis by approximately the same angle, 44° ± 2°. They produce a total translation of 115.89 Å along *c*. The tilt direction and displacement for individual duplexes result in a 6_1 screw pseudosymmetry along *c*.

The duplexes interact with each other through the guanines of two terminal base pairs forming N2−N3 H-bonds. The terminal O3′H group forms an additional hydrogen bond with the N2 atom of a third guanine in the next duplex.

In the crystal, the columns of H-bonded duplexes pack in a pseudohexagonal array with a separation of about 25 Å between the axes of the columns. A group of six consecutive duplexes has 12 pseudodyads approximately perpendicular to *c*, which correspond to the end and center of each duplex. The true crystal symmetry is $P2_12_12_1$, and neighbor duplexes along *c* (shown in Figure 1) have similar but not identical conformations. A group of three of them corresponds to the asymmetric unit, and the true intrinsic symmetry of the endless motif of H-bonded duplexes drops to 2_1 screw symmetry along *c*.

Dodecamer Packing. The dodecamer d(CGCTCTAGAGCG) is packed with interactions between ends similar to those found in d(CGCGAATTCGCG) and in many other dodecamers (Dickerson et al., 1987), as shown in Figure 4. A hydrogen bond between the terminal O3′H group and the N2 atom of the third guanine in the next duplex is also present, as found in the octamer. However, neighbor columns of duplexes are packed in a way different to that found by Dickerson et al. (1987): in our case the ends of molecules in neighbor columns are located at the same height, and a stacking interaction between terminal C13 cytosines of two duplexes is found, as shown in Figure 4. This feature was already reported in another dodecamer structure (DiGabriele & Steitz, 1993), which crystallized in a different cell, also with two duplexes in the asymmetric unit. A view of the packing arrangement in the unit cell has been presented elsewhere (Urpí et al., 1996).

A calcium ion is found which interacts with four phosphates of four different duplexes (Figure 2). It shows only four coordinations with approximate octahedral symmetry and an average oxygen−ion distance 2.23 Å. This calcium ion could be essential for crystallization of the dodecamers in the unit cell we have found.

General Structural Features. The conformational angles (α to ζ,χ) of the molecules studied have conventional values and show a variability similar to that reported for other oligonucleotide structures (Yanagi et al., 1991; Heinemann et al., 1994; Subirana et al., 1995). They are given as Supporting Information. The main chain angle δ, which defines the sugar conformation, is found in most cases in the usual C2′-endo region, with some exceptions which are different in each case. Although in the B form the C3′-endo conformation is more common in pyrimidines (Fratini et al., 1982; Subirana et al., 1995), it is also found in purines.

The three crystallographically different octamer duplexes have an almost identical structure, except in the G6−C7 step, where small differences are found. For example, the values of the twist parameter are 38.7°, 40.8°, and 36.6° in that step. The two dodecamer duplexes present more differences among themselves, as shown in Figures 1 and 5. The two duplexes can be superimposed with a rms = 0.79 Å.

The propeller twist shows striking variations. In the octamers it has an alternating pattern starting from the center, as is apparent from the values given in Table 2, and the molecular symmetry is rather well preserved. In the dodecamers the pattern of propeller twist (Table 3) is more complex and not symmetric. The large propeller twist found in some base pairs gives rise to three-center hydrogen bonds in the narrow groove of most of the CT/AG steps, which have already been described by Yanagi et al. (1991) in their comparative studies of several structures. The N2 of guanine forms a bifurcated hydrogen bond with the O2 of its Watson−Crick cytosine partner and with the O2 of the next thymine. Distances smaller than 3.5 Å between the N2 of guanine and O2 of thymine have been found in six of the eight CT/AG steps present in the dodecamers. In the decamer d(CTCTCGAGAG) recently analyzed by Goodsell et al. (1995) it can be shown that three-center hydrogen bonds are also present. It appears that this alternating pattern of three-center hydrogen bonds is a characteristic feature of duplexes containing $(CT/AG)_n$ sequences. An alternating twist pattern is also apparent, with the CT/AG step having a smaller twist than the TC/GA step.

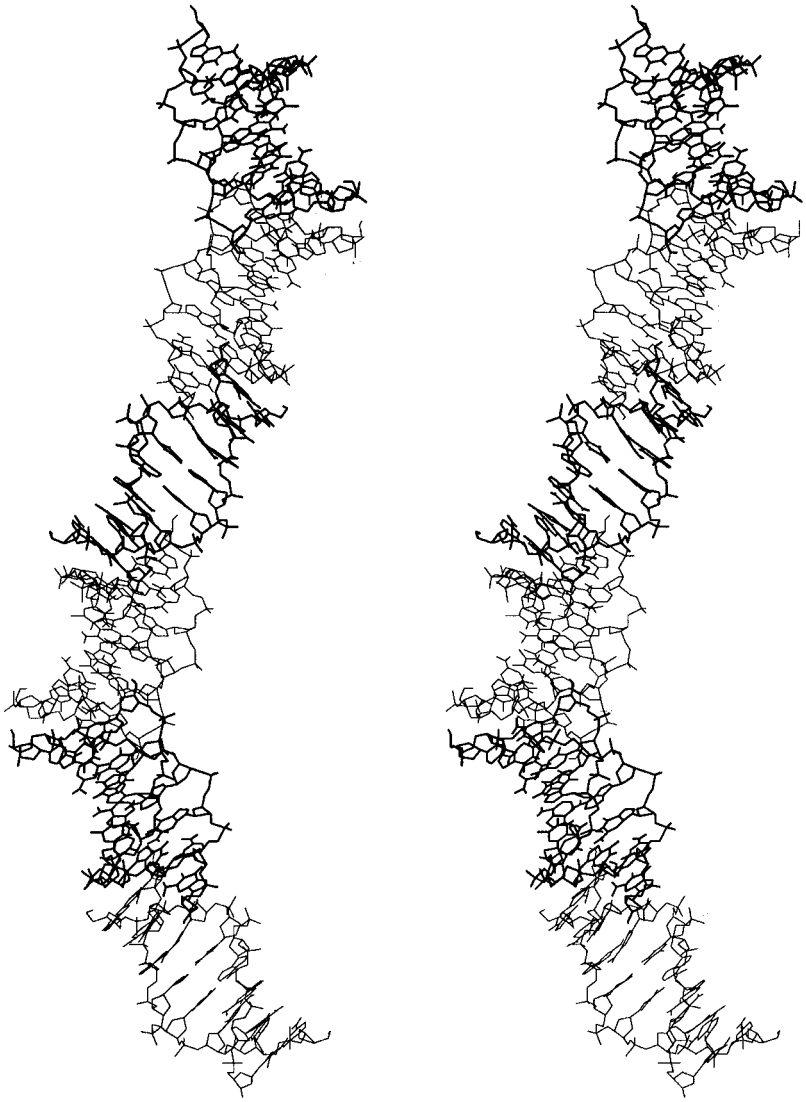

FIGURE 3: Stereo pair showing a column of six octamer duplexes which interact with each other through two terminal CG base pairs. The crystallographic *c* axis is vertical. Three consecutive octamer duplexes form the asymmetric unit of the crystal.

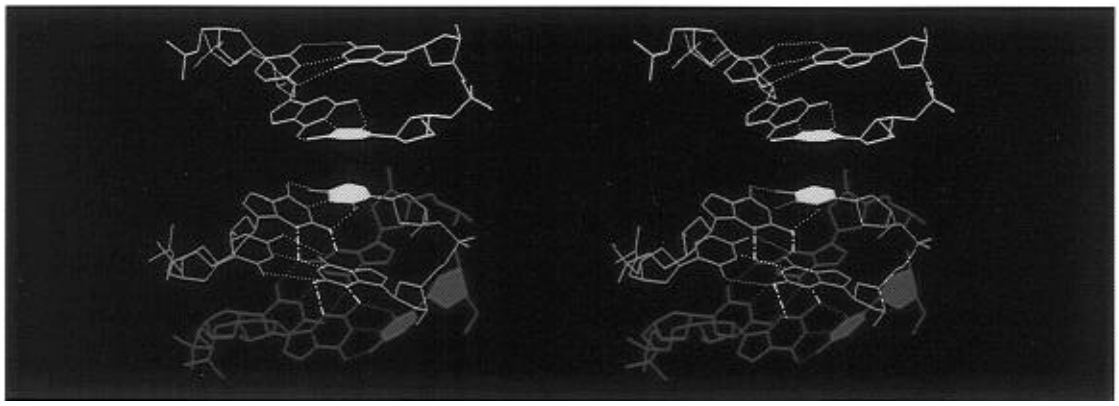

FIGURE 4: Detail of the interaction between the ends of three dodecamer duplexes. Each duplex is represented in a different color. The hydrogen bonds between guanines from different duplexes are shown as pink dashed lines. The stacking interaction between the terminal C13 cytosines (in white) of the different duplexes is clearly visible. As described in the text, cytosine 1 was disordered and did not appear in the electron density map. It is included in red in the drawing at its theoretical position. The yellow bases correspond to molecule D1, whereas the blue and red bases correspond to two D2 molecules related by a screw axis symmetry.

The rise and twist parameters are represented in Figure 5. The octamers are rather symmetric, whereas the two dodecamers show some differences at both ends which should be related to the different interactions with neighbor duplexes in the crystal. It is striking that the central CTAG sequence, which is identical in both molecules (octamer and dodecamer), has the same twist pattern in both cases but quite a different rise pattern. This may be related to the sudden increase (about 10°) in propeller twist of the dodecamers in the central TA step, whereas in the octamers propeller twist increases only 4° in the equivalent step. These observations illustrate the intrinsic variability of DNA sequences as a

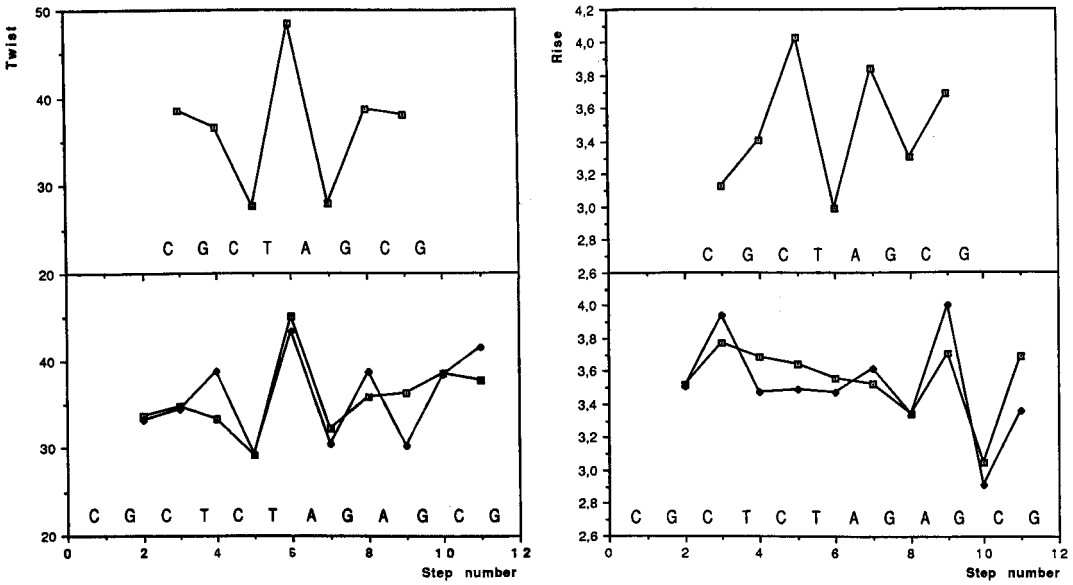

FIGURE 5: Comparison of twist and rise for the oligonucleotide duplexes indicated. The values for both dodecamers are given (□, dodecamer 1; ●, dodecamer 2). The first base pair is not included in the plots, since cytosine 1 could not be positioned, as described in the text.

function of environment, a question which has been discussed elsewhere (Urpí et al., 1996) and will be analyzed in more detail below.

Hydration Pattern. The resolution of our data (2.25 and 2.47 Å in either crystal) does not warrant a detailed study of hydration. Only some scattered water molecules which interact with the oligonucleotides could be asigned, but no consistent pattern of hydration could be detected. About half of the water molecules in contact with the oligonucleotides interact with the bases and the rest with the phosphodiester backbone. The three octamer duplexes respectively have six, seven, and nine associated water molecules, but none of them appears in the same relative position in the three duplexes. In the case of dodecamers a calcium ion and 47 water molecules were detected in the asymmetric unit. Among them, 26 were found to interact directly with the dodecamer duplexes, 4 of which formed bridges between two duplexes in the crystal. Only one of them was found in the same relative position in both dodecamer duplexes in the asymmetric unit, with a strong interaction (2.5; 2.7 Å in either duplex) with the O6 of guanine 10. Its position corresponds to one of the standard water molecules described by Schneider and Berman (1995).

It is interesting to note that no consistent pattern of hydration was found to be present in the major groove. In the *trp* repressor−operator complex it has been found (Otwinowski et al., 1988) that the protein/base interaction takes place through some water molecules, which are also present in a decamer structure with a related sequence (Shakked et al., 1994). Six clearly defined water molecules are found in the central CTAG sequence, as discussed in detail by Shakked et al. (1994). In the oligonucleotide structures we have studied, only one or two of these water molecules are found in any of the three octamer and two dodecamer duplexes. Furthermore, in each case different types of water molecules are involved. Although the resolution of our data does not allow a detailed description of hydration, it is clear that the striking hydration pattern of the *trp* repressor−operator complex is not found in the structures we have studied. This is not surprising, since the pattern of hydration depends very much on the conformational parameters of the oligonucleotides (Goodsell et al., 1994a; Figure 6), which vary significantly in each case, as it will be shown in detail in the next section.

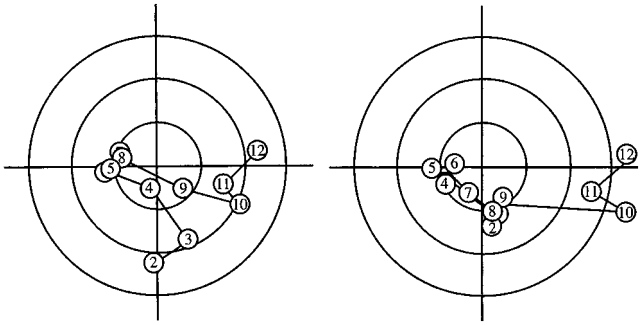

FIGURE 6: Normal vector plots of the two dodecamer helices in the asymmetric unit (Dickerson et al., 1994). According to this representation, molecule D1 (at left) is bent at both ends, while molecule D2 (at right) is only bent at one end. The numbers correspond to the base pairs as given in Table 3. The circles correspond to angles of 5°, 10°, and 15°.

Structure of the Central TA Step. In a previous publication (Urpí et al., 1996) we have compared in detail the conformational parameters of the central CTAG region in different sequences which have been crystallized either alone or combined with proteins. We found significant conformational differences in the various structures compared. In particular, the roll parameter of the TA step has a comparatively large positive value (12°) in the *trp* repressor−operator complex, while in the structures described here it is either small or negative (Tables 2 and 3 and Supporting Information). In the related TTAA sequence the TA step is also placed between a pyrimidine and a purine. Two different structures have been determined (Goodsell et al., 1994a) for oligonucleotides which contain this sequence. In that case the two conformations have different twist values (38.4° in the trigonal form, 31.1° in the orthorhombic form) but a similar roll angle of about 8°. From these observations it has been suggested (Goodsell et al., 1994a; Dickerson et al., 1994) that the TA step always has a positive roll, a conclusion which is not supported by the results we have found with the CTAG sequence: the TA step is also capable of changing its roll angle. These observations agree with the generally accepted fact that pyrimidine−purine steps have

a strong conformational variability, due in part to poor stacking of base pairs.

Curvature and Bending. The oligonucleotides in the crystal interact through their ends with neighbor molecules as shown in detail in Figures 3 and 4. Such interaction had already been found (Dickerson et al., 1987; DiGabriele & Steitz, 1993) in all previously crystallized dodecamers with a starting CGY sequence. Our work demonstrates that this interaction has some flexibility and may take place in a new space group ($P2_1$ in the dodecamer) and also in octamers. In fact, there is a considerable range of variability in the N2−N3 hydrogen bond distances in the guanine−guanine interactions, with 2.66 and 3.50 Å being the extreme values found in the oligonucleotide structures presented here.

There are also phosphate−phosphate interactions between neighbor molecules mediated by water and ions, most of which can only be detected as short phosphate−phosphate distances due to molecular disorder. In favorable cases, such as the one shown in Figure 2, the ions may occupy fixed positions and can be located in the crystal.

As a result of all these interactions and the intrinsic characteristics of the oligonucleotide base sequence, the molecules in the crystal are considerably distorted, as shown in Figure 1. It appears that dodecamers optimize inter- and intramolecular interactions in the crystal through bending and other types of distortion, while octamers may also change its inclination. The distorsions of the oligonucleotide molecules may be analyzed with the help of the roll parameter, as we have done in detail for the central CTAG sequence (Urpí et al., 1996). A graphic representation of the bends, which are mainly due to variations in roll, can also be obtained from the normal vector plots (Dickerson et al., 1994) as shown in Figure 6.

It should be noted that the three methods (roll values, normal vectors, and program CURVES) we have used to describe the distortions of the dodecamer molecule are related but not identical. Dickerson et al. (1996) have recently provided a detailed analysis of bending in oligonucleotide crystals by the normal vector plot and shown how it relates to the roll parameter. In this context we should note that the normal vector plot does not detect the curvature of the molecule when it is due to sliding of the bases. Some base steps may remain with their base pairs parallel (roll = 0°) but show a noticeable local curvature due to sliding, a feature which may be detected by the program CURVES.

The results presented in Figures 1 and 6, as well as in Table 3, show that the dodecamer we have studied can have two different stable conformations in the same crystal, which are those favored by the interactions present under these conditions. Since both chemically identical ends of each dodecamer show different curves, we have four patterns of curvature for the CGCTCT half-duplexes in the asymmetric unit. From such observations it is not possible to decide which would be the pattern of bending of this sequence in solution or upon interaction with proteins. It appears that crystal studies may only show which types of bending are possible (Goodsell et al., 1994b; Dickerson et al., 1996) but we cannot decide which is the most likely bending for a particular sequence. The central CTAG sequence also shows different patterns of bending as a function of its environment (Urpí et al., 1996).

Finally, we should note that the octamer molecule has very little distortion; the normal vector plot (not shown) does not detect any large bending as is also apparent from the comparatively small values of roll (Table 2). Only in the central TA step is an appreciable negative value of roll found, which appears to be much larger in solution (Manderville et al., 1995).

ACKNOWLEDGMENT

We are thankful to Drs. F. Azorín, C. Lawson, and S. Phillips for stimulating discussions and to Dr. R. Lavery for sending the program CURVES.

SUPPORTING INFORMATION AVAILABLE

Tables of conformational angles and structural parameters for the five independent molecules presented in this paper, calculated with NEWHELIX93 (5 pages). Ordering information is given on any current masthead page.

REFERENCES

Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., & Schneider, B. (1992) *Biophys. J. 63*, 751−759.

Brunger, A. T. (1992) *X-PLOR Manual*, Yale University, New Haven, CT.

Dickerson, R. E., Goodsell, D. S., Kopka, M. L., & Pjura, P. E. (1987) *J. Biomol. Struct. Dyn. 5*, 557−579.

Dickerson, R. E., Goodsell, D. S., & Neidle, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3579−3583.

Dickerson, R. E., Goodsell, D., & Kopka, M. L. (1996) *J. Mol. Biol. 256*, 108−125.

DiGrabriele, A. D., & Steitz, T. A. (1993) *J. Mol. Biol. 231*, 1024−1039.

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686−14707.

Goodsell, D. S., Kaczor-Grzeskowiak, M., & Dickerson, R. E. (1994a) *J. Mol. Biol. 239*, 79−96.

Goodsell, D. S., Grzeskowiak, K., Kopka, M. L., & Dickerson, R. E. (1994b) in *Structural Biology: The State of the Art* (Sarma, R. H., & Sarma, M. H., Eds.) pp 215−220, Adenine Press, Albany, NY.

Goddsell, D. S., Grzeskowiak, K., & Dickerson, R. E. (1995) *Biochemistry 34*, 1022−1029.

Hahn, M., & Heinemann, U. (1993) *Acta Crystallogr. D49*, 468−477.

Heinemann, U., Alings, C., & Hahn, M. (1994) *Biophys. Chem. 50*, 157−167.

Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67−71.

Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.

Manderville, R. A., Ellena, J. F., & Hecht, S. M. (1995) *J. Am. Chem. Soc. 117*, 7891−7903.

Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321−329.

Schneider, B., & Berman, H. M. (1995) *Biophys. J. 69*, 2661−2669.

Shakked, Z., Guzikevich-Guerstein, G., Frolow, F., Rabinovich, D., Joachimiak, A., & Sigler, P. B. (1994) *Nature 368*, 469−473.

Subirana, J. A., Salas, X., Urpí, L., Font, E., & Verdaguer, N. (1995) in *Nuevas Tendencias en Cristalografía* (*New Trends in Crystallography*) (Cano, F. H., Foces-Foces, C., & Martínez-Ripoll, M., Eds.) pp 223−240, Ediciones del Consejo Superior de Investigaciones Científicas, Madrid.

Urpí, L., Tereshko, V., Malinina, L., Huynh-Dinh, T., & Subirana, J. A. (1996) *Nat. Struct. Biol. 3*, 325−328.

Westhof, E. (1984) in *Handbook of Biochemistry, Nucleic Acids* (Fasman, G. D., Ed.) Vol. II, pp 411−422, CRC Press, Boca Raton, FL.

Westhof, E. (1987) *J. Biomol. Struct. Dyn. 5*, 581−599.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 287*, 755−758.

Yanagi, K., Privé, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201−214.

BI961048L